



PCT/AU98/01023  
09/581286

REC'D	12 JAN 1999
WIPO	PCT

Patent Office  
Canberra

5

I, KIM MARSHALL, MANAGER EXAMINATION SUPPORT AND SALES, hereby certify that the annexed is a true copy of the Provisional specification in connection with Application No. PP 3128 for a patent by CSL LIMITED filed on 23 April 1998.

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)



WITNESS my hand this Fifth  
day of January 1999

KIM MARSHALL  
MANAGER EXAMINATION SUPPORT AND  
SALES

**AUSTRALIA**

**Patents Act 1990**

**CSL LIMITED**

**PROVISIONAL SPECIFICATION**

*Invention Title:*

*Porphyromonas gingivalis antigens and probes*

The invention is described in the following statement:

*Porphyromonas gingivalis* antigens and probes

FIELD OF THE INVENTION

5           The present invention relates to *P. gingivalis* nucleotide sequences, *P. gingivalis* polypeptides and probes for detection of *P. gingivalis*.

BACKGROUND OF THE INVENTION

10           Periodontal diseases are bacterial-associated inflammatory diseases of the supporting tissues of the teeth and range from the relatively mild form of gingivitis, the non-specific, reversible inflammation of gingival tissue to the more aggressive forms of periodontitis which are characterised by the destruction of the tooth's supporting structures. Periodontitis is associated  
15 with a subgingival infection of a consortium of specific Gram-negative bacteria that leads to the destruction of the periodontium and is a major public health problem. One bacterium that has attracted considerable interest is *P. gingivalis* as the recovery of this microorganism from adult periodontitis lesions can be up to 50% of the subgingival anaerobically  
20 cultivable flora, whereas *P. gingivalis* is rarely recovered, and then in low numbers, from healthy sites. A proportional increase in the level of *P. gingivalis* in subgingival plaque has been associated with an increased severity of periodontitis and eradication of the microorganism from the cultivable subgingival microbial population is accompanied by resolution of  
25 the disease. The progression of periodontitis lesions in non-human primates has been demonstrated with the subgingival implantation of *P. gingivalis*. These findings in both animals and humans suggest a major role for *P. gingivalis* in the development of adult periodontitis.

*P. gingivalis* is a black-pigmented, anaerobic, asaccharolytic,  
30 proteolytic Gram-negative rod that obtains energy from the metabolism of specific amino acids. The microorganism has an absolute growth requirement for iron, preferentially in the form of haeme or its Fe(III) oxidation product haemin and when grown under conditions of excess haemin is highly virulent in experimental animals. A number of virulence  
35 factors have been implicated in the pathogenicity of *P. gingivalis* including the capsule, adhesins, cytotoxins and extracellular hydrolytic enzymes.

In order to develop an efficacious and safe vaccine to prevent *P. gingivalis* colonisation it is necessary to identify and produce antigens that are involved in virulence that have utility as immunogens to generate neutralising antibodies. Whilst it is possible to attempt to isolate antigens directly from cultures of *P. gingivalis* this is often difficult. For example as mentioned above, *P. gingivalis* is a strict anaerobe and can be difficult to isolate and grow. It is also known that, for a number of organisms, when cultured *in vitro* that many virulence genes are down regulated and the encoded proteins are no longer expressed. If conventional chemistry techniques were applied to purify vaccine candidates potentially important (protective) molecules may not be identified. With DNA sequencing, as the gene is present (but not transcribed) even when the organism is grown *in vitro* it can be identified, cloned and produced as a recombinant DNA protein. Similarly, a protective antigen or therapeutic target may be transiently expressed by the organism *in vitro* or produced in low levels making the identification of these molecules extremely difficult by conventional methods.

With serological identification of therapeutic targets one is limited to those responses which are detectable using standard methods such as Western Blotting or ELISA. The limitation here is the both the level of response that is generated by the animal or human and determining whether this response is protective, damaging or irrelevant. No such limitation is present with a sequencing approach to the identification of potential therapeutic or prophylactic targets.

It is also well known that *P. gingivalis* produces a range of broadly active proteases (University of Melbourne International Patent Application No PCT /AU 96/00673, US Patent Nos 5,475,097 and 5,523,390), which make the identification of intact proteins difficult because of their degradation by these proteases.

## SUMMARY OF THE INVENTION

The present inventors have attempted to isolate *P. gingivalis* nucleotide sequences which can be used for recombinant production of *P. gingivalis* polypeptides and to develop nucleotide probes specific for *P. gingivalis*. The DNA sequences listed below have been selected from a

large number of *P. gingivalis* sequences according to their indicative potential as vaccine candidates. This intuitive step involved comparison of the deduced protein sequence from the *P. gingivalis* DNA sequences to the known protein sequence databases. Some of the characteristics used to  
 5 select useful vaccine candidates include; the expected cellular location, such as outer membrane proteins or secreted proteins, particular functional activities of similar proteins such as those with an enzymatic or proteolytic activity, proteins involved in essential metabolic pathways that when inactivated or blocked may be deleterious or lethal to the organism, proteins  
 10 that might be expected to play a role in the pathogenesis of the organism eg. red cell lysis, cell agglutination or cell receptors and proteins which are paralogues to proteins with proven vaccine efficacy. DNA sequences that were considered to be poor vaccine candidates and not selected include those that code for proteins involved in replication, non-essential proteins  
 15 involved in cellular processes and those proteins present at sites that would be unlikely to be affected by immune mediators such as those found in the bacterial cytoplasm or inner membranes.

In a first aspect the present invention consists in an isolated  
*P. gingivalis* nucleotide sequence, the nucleotide sequence consisting of or  
 20 including a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9 and sequences complementary thereto.

In a second aspect the present invention consists in an isolated  
 25 *P. gingivalis* polypeptide, the polypeptide being encoded by a nucleotide consisting of or including a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8 and SEQ ID NO: 9.

In a third aspect the present invention consists in a nucleotide probe  
 30 specific for *P. gingivalis*, the probe including a detectable label and a nucleotide sequence of at least 15 nucleotides, the nucleotide sequence being derived from a sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, and sequences  
 35 complementary thereto.

In a fourth aspect the present invention consists in an isolated *P. gingivalis* polypeptide, the polypeptide consisting of or including a sequence selected from the group consisting of SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27.

In a fifth aspect the present invention consists in an isolated polynucleotide, the polynucleotide encoding a polypeptide consisting of or including a sequence selected from the group consisting of SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27.

In a sixth aspect the present invention consists in a composition for use in raising an immune response in an animal directed against *P. gingivalis* the composition including an acceptable carrier and/or adjuvant and at least one polypeptide, the at least one polypeptide having a sequence selected from the group consisting of SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27.

Throughout this specification, unless the context requires otherwise, the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element or integer or group of elements or integers but not the exclusion of any other element or integer or group of elements or integers.

## DETAILED DESCRIPTION

Preparation of the *P. gingivalis* library for sequencing.

To determine the DNA sequence of *P. gingivalis* genomic DNA was isolated from *P. gingivalis* strain W50 (ATCC 53978) essentially by the method

described by Mamur J. (1961)(1). Cloning of DNA fragments was performed essentially as described by Fleischmann *et al.*, (1995)(2). Briefly, purified genomic DNA from *P. gingivalis* was nebulized to fragment the DNA and was treated with Bal31 nuclease to create blunt ends then run twice on  
 5 preparative 1% agarose gels. DNA fragments of 1.6-2.0 kb were excised from the gel and the DNA recovered. This DNA was then ligated to the vector pUC18 (*Sma*I digested and dephosphorylated; Pharmacia) and electrophoresed on a 1% agarose preparative gel. The fragment comprising linear vector plus one insert was excised, purified and this process repeated  
 10 to reduce any vector without insert contamination. The recovered vector plus insert DNA was blunt-ended with T4 DNA polymerase, then a final ligation to produce circular DNA was performed. Aliquots of Epicurian Coli Electroporation-Competent Cells (Stratagene) were transformed with the library DNA and plated out on SOB agar antibiotic diffusion plates  
 15 containing X-gal and incubated at 37°C overnight. Colonies with inserts appeared white and those without inserts (vector alone) appeared blue. Plates were stored at 4°C until the white clones were picked and expanded for the extraction of plasmid DNA for sequencing.

## 20 DNA sequencing

Plasmid DNA was prepared by picking bacterial colonies into 1.5ml of LB, TB or SOB broth supplemented with 50-100ug/ml Ampicillin in 96 deep well plates. Plasmid DNA was isolated using the QIAprep Spin or QIAprep  
 25 96 Turbo miniprep kits (QIAGEN GmbH, Germany). DNA was eluted into a 96 well gridded array and stored at -20C.

Sequencing reactions were performed using ABI PRISM Dye Terminator and ABI PRISM BIGDye Terminator Cycle Sequencing Ready Reaction kits with AmpliTaq DNA polymerase FS (PE Applied Biosystems, Foster City, CA) using the M13 Universal forward and reverse sequencing  
 30 primers. Sequence reactions were conducted on either a Perkin-Elmer GeneAmp 9700 (PE Applied Biosystems) or Hybaid PCR Express (Hybaid, UK) thermal cyclers. Sequencing reactions were analysed on ABI PRISM 377 DNA sequencers (PE Applied Biosystems).

The sequences obtained are set out below.

35 SeqID#1-9 represents the DNA sequence that encodes proteins in SeqID#10-27. Proteins in SeqID#10-18 are the entire open reading frame

from DNA SeqID#1-9. Proteins in SeqID#19-27 are the proteins encoded by DNA SeqID#1-9 from their putative initiation codon. The initiation codon was calculated from sequence homology alignment using FastX or by the ORF prediction program GeneMark.

5       As will be understood by those skilled in the art open reading frames (ORFs) may be readily identified. ORFs may be determined using two methods, for example, alignments from FastX search results may be used to define the start and end positions of coding regions if sufficient protein homology is present. Alternatively, protein coding regions may be identified  
10       using the ORF identification program GeneMark (3) using a training matrix based on published *P.gingivalis* sequences. This matrix may be further refined by adding ORFs identified from the results of homology searching and ORFs identified by GeneMark. The program PSORT (4) may be used for the detection of signal sequences and the prediction of protein cell  
15       localisation. A UNIX version of TopPred (5) may be used to identify potential membrane spanning domains.

#### **DNA sequence analysis**

20       Raw trace data files from the ABI 377 sequencer were manually trimmed using Staden Pregap (Laboratory of Molecular Biology, Medical Research Council, UK) running on a Sun Microsystem computer. Trimmed files were assembled into contigs using Staden Gap v4.1 and exported as a consensus file in FastA format. This consensus was converted into GCG  
25       format files and analysed for homology using the FASTX algorithm on a non-redundant protein database compiled by ANGIS (Australian Genomic Information Service, University of Sydney). Individual FASTX search results were examined for significant homology by statistical probability and amino acid alignments.

The results are set out in Table 1.



It will be appreciated by persons skilled in the art that numerous variations and/or modifications may be made to the invention as shown in the specific embodiments without departing from the spirit or scope of the invention as broadly described. The present embodiments are, therefore, to  
5 be considered in all respects as illustrative and not restrictive.

Dated this 23rd day of April 1998

CSL LIMITED

Patent Attorneys for the Applicant:

F.B. RICE & CO.

**References.**

1. Mamur, J. (1961) A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* 3, 208-218.
2. Fleishmann, R.D. et al. (1995) Whole genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496-512.
3. Borodovsky M, Rudd KE, and EV Koonin. (1994). Intrinsic and extrinsic approaches for detecting genes in a bacterial genome. *Nucleic Acids Res.* 22:4756-4767.
4. Horton, P. and Nakai, K. (1996). A probabilistic classification system for predicting the cellular localization sites of proteins. *Intellig. Syst. Mol. Biol.* 4: 109-115.
5. Claros MG and G von Heijne. (1994). TopPred II: an improved software for membrane protein structure predictions. *Comput. Appl. Biosci.* 10: 685-686.

Table 1

SeqID#	Length of protein in SeqID		FastX homology results			
		Homology description	Length of protein homolog	% identity	Overlap	E value
1, 10, 19	195aa	Macrophage infectivity potentiator, <i>Legionella oakridgensis</i>	234aa	50	201aa	3.40E-36
2, 11, 20	271aa	Outer membrane lipoprotein P4, <i>Haemophilus influenzae</i>	274aa	36	254aa	8.20E-27
3, 12, 21	296aa	Heme uptake protein B, <i>Bacteroides fragilis</i>	160aa	71	153aa	3.20E-48
4, 13, 22	981aa	Protein-export membrane protein secD, <i>Helicobacter pylori</i>	503aa	44	274aa	5.20E-43
5, 14, 23	76aa	Alpha-hemolysin, <i>Aeromonas hydrophila</i>	85aa	57	67aa	1.70E-14
6, 15, 24	417aa	Hemolysin TYLC, <i>Serpulina hyodysenteriae</i>	268aa	31	235aa	9.10E-23
7, 16, 25	672aa	Outer membrane protein A, <i>Escherichia fergusonii</i>	243aa	48	88aa	2.80E-11
8, 17, 26	324aa	Hemolysin A, <i>Prevotella melaninogenica</i>	332aa	61	306aa	1.60E-86
9, 18, 27	313aa	Hemagglutinin, <i>Prevotella intermedia</i>	309aa	43	305aa	2.80E-45

## (2) INFORMATION FOR SEQ ID NO:1

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 603 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...603

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1

```

gaactaagca agtgttacat ggataaagtg agctatgctc tgggattgag catcggtaat      60
aatttcaagt cttcgggcat cgacagcgctc gttatggatg atttcatgca aggtctgtct      120
gatgtactgg aagaaaaagc ccctcagctc tcgtatgacg aggccaagcg cgaaatagag      180
gcgtatttca tggatttgca gcagaaggct gtcaaaactga acaaagaggc cggagaagaa      240
ttcctcaaga taaatgcaca caaggaagggt gtgacgacct taccgagcgg cttgcaatac      300
gaagtcatta agatgggaga gggcccgaaa ccaccccttt cggacacggt aacctgtcat      360
tatcacggta cgctcatcaa cggtatcggt ttcgatagct ctatggacag gggagaaccg      420
gccagtttcc ctctaagagg agttatagcc ggctggacgg agattcttca attaatgcct      480
gtaggatcca agtggaagt aactataccg agcgatctgg cgtatggaga tcgtgggtgcc      540
ggcgaacata tcaaaccggg tagtacgctc atttttataa tcgaattatt gagtatcaac      600
aaa

```

## (2) INFORMATION FOR SEQ ID NO:2

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1119 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...1119

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2

```

gggtcttggt gaagtagccc agcaattggc cgatcagggt gttcgcgttg tgatcgccgg      60
attggacatg gactttcgac gtcagccttt cggacctatg ccgggcttgt gtgccatagc      120
cgactccgtg accaaagttc atgccgtgtg tgtggaatgc ggccgattgg ccagctattc      180
tttccgtcgt gtccaaggcg atcagcaagt gatgctgggc gaactgaacg aatacagtcc      240

```

cctctgcaga	acctgctaca	ggaaatgcag	ttctccccc	caaacagaag	aaatccattc	300
gacaatatga	atagcagaca	tctgacaatc	acaatcattg	cgggcctctc	cctcttttga	360
ctgacattgg	gcggctgctc	cgtagcccaa	caagatacgc	agtggactct	cggcggaag	420
ctctttactt	cggcggtgat	acaacgttcg	gccgaatata	aagcgctttg	cattcaggca	480
tacaacatcg	ctacggaaag	agtggacgct	ctaccggcag	aacgtaaaca	aggagatagg	540
ccttatgcc	tcgtaacgga	catagacgaa	accatttttg	acaatacgcc	taactccgtg	600
tatcaggctc	tcaggggcaa	ggattatgat	gaagagactt	gggggaaatg	gtgtgcacag	660
gccgatgccg	acacactggc	aggagctttg	tctttcttcc	tccatgcagc	gaacaagggg	720
atcgaggctc	tttacgtcac	caaccgcaga	gacaatctgc	gcgaagcaac	tcttcagaac	780
cttcagcggt	acggattccc	ctttgccgat	gaagaacatt	tgcttacgac	ccatgggcca	840
tccgacaaag	aaccccgctc	gctcaaaata	caagaacagt	atgaaatagt	attgctcata	900
ggagacaact	tgggcgactt	ccaccacttc	ttcaatacga	aagaagagtc	cggacgcaaa	960
caggctctgg	gcctgacagc	cggggagttt	ggccggcact	tcacatgct	gccaatccc	1020
aactacggat	cttggaacc	ggcatggtac	ggcggaagt	atccgccact	gcccgaaga	1080
gacaaagcac	ttaaacaact	gcactcacag	aacagcaga			1119

## (2) INFORMATION FOR SEQ ID NO:3

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 933 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...933

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3

ccatataatg	tccaatctat	tagcaacaag	acgattaaaa	aacaaatgga	aaacttaaag	60
aacattcagc	ccagagagga	tttcaactgg	gaagagtttg	aggccggtgg	cgtccatgct	120
gccgtgagtc	gtcaggagca	ggaagctgct	tatgacaaaa	cgctcaatac	catcaaggaa	180
aaggaagtgg	taatgggtag	ggtaactgct	atcaacaagc	gtgaagtggg	tatcaatgta	240
gggtacaaat	cggaaggtgt	ggtacctgca	acagaattcc	gctacaatcc	cgaactcaaa	300
gtgggagacg	aagtggaaat	ttatatcgag	aatcaggaag	ataagaaggg	ccagctcgtc	360
ttgtctcacc	gcaaggctcg	tgccgctcgc	tcttgggagc	gcgtgaacga	ggctctcgaa	420
aaagacgaaa	tcgtaaaggg	ctatgtgaag	tgctcgtaac	agggtgggat	gatcgctgat	480
gtattcggt	tcgaggcttt	cctcccggga	tcacagatcg	acgtgcgccc	cattcgcgac	540
tacgatgcat	tcgttgagaa	gacgatggag	ttcaagattg	tgaaaatcaa	tcaagaatat	600
aagaatgtag	ttgtttccca	caagggtgctc	atcgaagcag	agctcgaaca	acagaagaaa	660
gaaatcatcg	gcaagctcga	aaaaggggcag	gtactcgaag	gtatcgtaaa	gaatattact	720
tcttacggag	tatttatcga	cctcggtgga	gtggatggtc	ttatccatat	cactgacctt	780
tcattggggtc	gtgtgggtca	tccggaagaa	atcgtaacgc	tggatcagaa	gatcaatgtc	840
gttatcctcg	actttgatga	agatcgcaag	cgtatcgctc	tcggactcaa	acagctgatg	900
ccctcactct	tgggatgctc	tcgacagcga	gct			933

## (2) INFORMATION FOR SEQ ID NO:4

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2955 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...2955

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4

gtaacagaca	aaatgcaaaa	caaaggattt	gtgattgtta	tcacatcggc	tctggccatc	60
atctgtgcgt	tttacctgtc	attctctttc	gttacgaacc	gttacgaaaa	gaaggctaag	120
gcgatgggcg	atggtgcccg	aatggcctat	cttgattcca	tgtcgaatga	gaaggctctg	180
ttcggctaca	cgctgaaaga	agctcaagcc	cagcaaattg	gtccttggcct	tgacttaaag	240
gggggtatga	acgttatctt	gaaacttaac	gcaagcgatc	tgcttcgtaa	cctctctaac	300
aaaagtttgg	atcccaactt	caacaaagct	ctggagaatg	ctgccaagag	cacggagcaa	360
tccgacttca	tcgatatttt	cgtgaaggaa	tatcgcaagc	tcgatcccaa	cggtcgcttg	420
gccgttatct	tcggttcggg	tgaccttcgc	gaccagatta	ccgcaaagtc	tacggatgca	480
gacgtagtgc	gtctgtctca	agaaaaatat	aatagtgtct	tagaagcttc	gttcaatgtg	540
ctccgtgtct	gtatcgatgc	tttcgggtgt	gttgcaacct	atttgcagcg	attggaaggg	600
caagggcgta	tccttgctga	actccccgga	gtgaaagacc	ctgagcgtgt	tcgtaccctt	660
ttgcaacgca	gtgccaacct	acagttctgg	cgtacataca	aattcgaaga	ggtcagcgga	720
gacttgatcg	ctgccaatga	tcgtctgagc	ggaattggcta	tgaacaacac	ggatgctacc	780
ccggaacag	agcctgcaac	tactgactct	gtagctgcaa	cagccgattc	tgctgtctga	840
caagctgtag	ctgattctgc	tactgtagca	caaaaagagg	ccaaggatgc	tactcgtaaa	900
gacgcactct	tctctctgct	tactcccgtg	aatcggtggc	gtgcagtagt	gggtgtggct	960
cgtcgtgtca	atatggctca	gatatctgaa	atgctccagc	aagctcacga	tctgaagggt	1020
acacgtgaag	atgtgctttt	cctctggggg	gctaaagcaa	tcgaagacct	cgaaaccaaa	1080
aaggagaccg	acctctacga	actctatgct	attcgtacca	atcgtacggg	agatcctgat	1140
ttgggagggt	atgtagtgc	ttccgccaa	agtgatatcc	aaaatgactt	cggtcggttc	1200
gaaccgatcg	tttcgatgac	gatgaatgaa	gaagggtgtc	gtaaatgggc	gcgtatcaca	1260
aaggataacg	tgggacgggc	aatcgctatc	gttttggtatg	gtgtggttta	ttctgtctcc	1320
aacgtgaatg	atgagatcac	gggcgggtcg	tctcagatct	ccgggcactt	caccgtggag	1380
gaggccgggt	accttgccaa	cgtactcaac	tccggtaaaa	tggtatgtac	ggtaagcatc	1440
gaacaggaaa	acgtgattgg	tcctacgtgt	gggtccgagt	ccattaaagc	cggattcttg	1500
tcgttcctgc	tcgctttggt	tatcctgatg	tgttacatgt	gtctggctta	cggtttcttg	1560
ccgggtctta	tcgcaaacgg	cgcattgatt	gtaaacagct	tcttcacatt	gggcgtattg	1620
gcttctttcc	atgccgtgct	gacctctcgc	gggtatcgag	gtttggtgct	gacgctgggt	1680
atggctgtgg	atgccaacgt	acttatcttc	gagcgtatca	aagaagagct	tcgtgccggg	1740
aagactccga	ttcgtgccgt	tacggatggt	tatggcaacg	ctttctctgc	catcttcgac	1800
tcgaacgtta	cgactattat	taccggtatc	atcctatttc	tctacgggac	ggggccgatt	1860
cgcgggtttg	ccactacgtt	gattatcggt	cttatcgctt	ctttcattac	ggctgtcttc	1920
ttgactcgta	tcgtcttcga	gaaactggcg	aaaaaagggtc	gtttggataa	gattacattc	1980
actacgagca	ttactcgcaa	tctccttgct	aatccctcat	acaacatctt	gggtaagcgc	2040
aagaccggct	ttatcattcc	ggtgattatc	atcgttttgg	gacttatagc	ttcatttaca	2100
atcgggtctca	ataggggtat	tgaattctcc	ggaggacgta	actacgtagt	taaattcgac	2160
cagcctgtat	cttccgaagc	cgttcgttcg	gccttgtctt	ctccccctgca	ggaaaaggta	2220
ttggttacct	ccatcggtac	tgaagggaca	gaggtgcgta	tatctacgaa	ctataagatc	2280
caggaggaaa	gcgaagaaac	tgaagcagag	attactgaca	aattgtatca	gagcctgaaa	2340
ggtttctaca	cccagcagcc	tactgtctgat	cagttcttgg	acaatatcat	tagctctcag	2400
aaagtaagtc	ccagtatgtc	gagtgcacatc	acgagaggtg	ctatttgggc	tgtgctgtta	2460
tcgatgatct	tcattggccat	ttatatctctg	attcgtcttc	gtgacatttc	tttctctgcc	2520
gggttattcg	tatctgtggc	cgctactaca	ttctgcatta	ttgctctgta	tcggttgctg	2580
tggaagattc	tgcccttcac	catggagatc	gatcagaact	tcacgcctgc	tattctggct	2640
atcatcggtt	actcgtctca	tgacaccgtg	gttgattttg	accgtatccg	agagacgatg	2700
aaattgtacc	ccaacagaga	tcgctatcag	gtgatcaacg	atgcccttaa	ttcaacattg	2760
ggtcgaacat	taaatacgtc	tttgactacg	tttatcgcta	tggttggaat	cttcactctt	2820

ggaggtgcta cgatgcgtag tttcacgttc tcgatcctgc tcggtatcgt taticggtaca	2880
tactctacgc tctttgttgc tacacccctt gcctacgaga tccaaaagcg caagctcaac	2940
aaagcagcta agaaa	2955

## (2) INFORMATION FOR SEQ ID NO:5

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 318 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...318

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5

ctttcgtaca gtggagagag cgatgcaaaa gagtctgatc agaattgccg gaaatgtacc	60
ttcatcggct ttgaaaaacg agtaaatcgc atgcgactga tcaaggcttt tctcgtgcaa	120
ctcttactgc tccccatttt cttctacaag cggttttatat cgccgcttac accgccttca	180
tgccggttta cccctcatg ttgcctctat gccatcgaag ccttacgtaa atatggcccg	240
ggcaaaggac tattgctgag catcaagcgt attctccgct gtcacccgtg ggggtggaagt	300
ggctatgacc ccgttccg	318

## (2) INFORMATION FOR SEQ ID NO:6

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1368 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

(ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...1368

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6

agaggaggac aaatccgccg acaccatacc gattcgtcaa ggggatcggg cagcaaagcc	60
tgccgcttct cctttcatgc cgagaccatc ggtttctcca accaccaag agcaaaaatg	120
aattacctgt acatactgat tacactttta ctctccggct tttttccgg tgctgagatt	180
gctttccttt cttcggacaa actgcgtctt gagttggaca ggaatagagg cgatctcaca	240
ggcagagcgt taaacttgct gtatcgacat ccggaccaac tggtgactac tctccttggt	300

ggtaataata	tcggttttgg	agtctatggt	ctgctgatgg	cgggattgct	ggccgcacct	360
ttggcgcaat	ggattgataa	cgatgctatg	atcgtcgttc	tccaatctgt	cttatccact	420
atcatcatac	tgtttaccgg	ggaatttcta	cccaaagcca	ttttcaagac	caatgccaat	480
atgatgatga	gggtattcgc	cctccctatc	gtagcgatct	attatctgct	ttatcctctg	540
tctaaactct	tcaccggttt	atctcgctct	tttattcgtc	tggtggacaa	gaattatgtg	600
cctacaacag	taggggttgg	gcgcgtagat	ctcgatcatt	atttggcaga	aaatatgtcc	660
ggagaaaacg	aacagaacga	cttgactacc	gaagtgaaaa	tcattccagaa	tgcgctggat	720
ttttccggtg	ttcaggtgcg	agactgcatg	atcccacgca	atgagatgat	agcatgtgag	780
ttgcaaacgg	atattgaagt	actcaaaacg	acttttatcg	ataccggttt	gtccaagatc	840
attatctaca	gacagaacat	agatgacgta	gtaggatata	tccattcgag	cgaaatgttt	900
cgtagggcaag	actggcaaaa	acgtatcaat	actactgtat	tcgtaccgga	aagcatgtat	960
gccaaataaac	tgatgcgact	actcatgcag	cgcaagaaaa	gcattgcgat	cgtcatcgat	1020
gaacttggag	gtacggccgg	aatgggtcaca	ttagaggatt	tggtagaaga	gattttcggg	1080
gacattgagg	acgaacacga	cactcgcaag	atcatagcca	aacagctcgg	ccctcatacc	1140
tatctgggtc	gtggtcgtat	ggaaatagat	gatgtgaacg	aacgttttgg	gttgctcctt	1200
cctgagctcg	acgactacct	taccgtggcc	ggatttatcc	tgaatagcca	tcaaaatata	1260
ccacaggcca	atgaggtcgt	ggagattgct	ccttatactt	ttaccattct	cagatcttct	1320
tccaccaaga	tcgaactggg	gaaaatgtcc	atcgacgacc	aatcgaac		1368

## (2) INFORMATION FOR SEQ ID NO:7

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2037 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: UNKNOWN

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...2037

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7

cctatccctt	ttctttctta	tatgtacagc	ggacatcata	aaatccatta	tccttttctt	60
atcctgttgg	tatgccttgc	ttttgctgcc	tgcaagagcg	tgaagttgaa	agatgcggag	120
aaggcacatg	atcgccaaga	gtataccaag	gctgccgata	tgtacaatac	attatacagg	180
cgtaccgcag	gaaagcaggt	ggagatgaaa	gcttatacgg	ctttccgatc	cggtgaaaac	240
tatcgtgccg	ccggcagaca	agccaaagct	ttgcgtggct	atctgaatgc	cagacgctac	300
gggtatccgg	attctgtggt	actgctccgt	ttggcacaga	cttatcagca	aggaggtaac	360
tataaggaag	ccgaggtact	cttccgtgga	tatctggaag	cttatccgaa	aagttatttt	420
gcagctatcg	gtttggaggg	gtgtctcttt	gcccgccagc	aaaaggaata	tcctacacgt	480
taccggatac	ggcgagctgc	cgagtggaat	tcggcacggg	gcgacttcgg	ccgggcctat	540
gcacccgatg	cttcggctct	ctatttcaca	tcgagcagaa	gcaaagacga	cggtttggat	600
aatagcagca	taacgggact	gaaacccaac	gacatttata	tcatacaaag	agatgcacaa	660
ggacgatggg	gacgtcccga	tagcgtgttc	ggaggaaatc	acactccatg	ggatgaaggc	720
gtgccaacga	tcacgcccga	tggtagtacc	atatattata	cgttggcgca	gcaaggagcc	780
gattacgacc	gtacggtaca	gatctattcc	gccgctcgga	gcggagaagg	cggttggagc	840
aacggttcgc	tcgtggacat	tatgcgcgat	tcgctccgta	tggtctctca	tcctctatg	900
tcggcatccg	gcgattacct	gtatttcgtc	agcaatatag	gcggtagcta	tgccggcaag	960
gatatttata	gtgtcaaggt	gtcggatcgt	tcttatgggt	caccggagaa	tttggggcct	1020
gatatacata	gcgcggggga	cgaaatgttt	cccttcatag	atggggatag	tacccttttc	1080
ttcgcttcgg	acgggacgcg	cggtctggga	ggactggata	ttttcaaagc	cacgtggagc	1140
tctaccggcc	aatggcatgt	agtcaatatg	ggacaaccgg	tcaattcctc	tgccgatgat	1200
ttcggtcttg	ctgtggagcc	taaaggcaaa	aacaaagaag	aagctttgcc	ggacaacgga	1260



gtcaaagg	tattttgt	caaccgag	gatgcacg	gatggccg	cctcttcc	1320
ttcgaact	cggctatc	caccgagat	caaggttat	tgatggac	agaagaaa	1380
cccatagc	gagccact	caggatcgt	ggcgaacg	gccccgt	acagggat	1440
gtgactac	gtgacgat	ctcctata	atgagcgt	agggcgat	tcgctatg	1500
atgcttg	gagcatcg	ttatttga	cagtacgt	aactcaag	cgataccg	1560
aagcagag	agacctac	tgtggact	ttccttgc	cgcgtaga	agccgagg	1620
ttgcaaaa	ttttctat	tttcgata	gctactct	gccccgaa	catgaaga	1680
ttggacga	tgattcgt	cctcacgg	aatccgga	ttcggatc	attgggtt	1740
catgccga	ggaaaggc	cgatgctt	aacctcgg	tatctgac	cagagccaa	1800
tccgtggt	attacctc	gagtcgtg	atagcggc	acaggctt	gtggaaag	1860
tacggtaa	ctgtccca	gacggtag	gccaaaat	ccgaacgg	cgatttcct	1920
aagggaag	atgtgtca	cgaggaat	gtagcacct	tgaccgag	gcagcagt	1980
gtctgcga	aactgaac	tcgtaccg	ttccgtgt	tcgaaga	ggttcgt	2037

## (2) INFORMATION FOR SEQ ID NO:8

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1011 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

## (ii) MOLECULE TYPE: DNA (genomic)

## (iii) HYPOTHETICAL: NO

## (iv) ANTI-SENSE: UNKNOWN

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Porphyromonas gingivalis

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...1011

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8

cctcagcc	tcggcctt	agaaataa	attaaacca	tgtgcctc	acccataa	60
gctccgat	catccgat	gctcgagc	gagctgac	ccgatcgt	tctgcgat	120
acaaacaa	ccggcaat	gatctatg	tttacggc	aagaagct	gcattgat	180
aaagaagt	gccgactg	agaagaag	tttcggca	atggcggg	tactggca	240
gcgatcga	tagacgat	cgacaccat	cccgaggc	acaaacag	gatcgtat	300
gatccgca	acaaggct	actcggag	taccgctt	tctatggg	ggacgttg	360
ttcgatac	atggcaag	tttgctgg	acggcaga	tgtttcgt	cagtgatg	420
tttttgc	attatctc	ctacacag	gaattggg	gttcgttc	gtcgtccg	480
taccaatc	cacggatg	cacaaagg	atttttgt	tggaacat	ttgggacg	540
atcggagc	tcactgtg	caatccag	gcactctt	tctatggc	ggtgaccg	600
tacaaagc	atgatcgg	agctcgca	ctgatcct	attttctc	caagcact	660
tccgatcc	aaggcttg	caagcctt	catcccct	cgatagag	cagtgcgg	720
gacgaagc	tgttctct	atccgact	gacacca	acaagact	caatatag	780
gtgcgca	tggttatc	tatccctc	ctcgtgag	catatat	tttgtctc	840
gagatgc	ttttcggc	tgcatgaa	gagtcctt	gagaggtg	ggaaaccg	900
atattcat	ctgtgggt	gatcctgg	gagaaaaa	aacggcac	agagagct	960
atcctcag	ggaacgaa	aaaaggtc	gacagtag	atggccga	a	1011

## (2) INFORMATION FOR SEQ ID NO:9

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 972 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: circular

- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: UNKNOWN
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Porphyromonas gingivalis*
- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION 1...972
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9

```

ctaacgttgt tttgttgcaa ctatttcaaa cagatgagag caaacatttg gcagatactt      60
tccgttttcgg ttctcttttt cttcgggaca gcgatcggac aggctcagag tcgaaaccgt      120
acatacgagg cttatgtgaa acagtacgcc gacgaagcta tccgacagat gagccgctac      180
aatataccgg caagcatcac catagcacag gctttgggtg agacaggagc cggagccagt      240
acactggcca gcgtacacaa caatcacttc gggatcaa at gccacaaatc gtggacgggc      300
aagcgcacct atcgtaccga cgatgcgccg aacgaatgct tccgcagcta ttcggccgct      360
cgcgaatcgt atgaagatca ttcccgattt ctgctccaac cagctatcg tcccctgttc      420
aaactcgaca gagaagacta tcggggctgg gctacggggt tgcaacgctg tggctatgcc      480
accaatcggg gctatgccaa tctgctgata aagatgggtg agctgtatga gctatatgct      540
ttggatcgcg agaagtaccc ctcatggttc cacaagtctt accccggggtc caacaaaaaa      600
tcccatcaaa cgaccaagca gaagcagagc ggactcaagc acgaagctta cttcagctac      660
ggactgctct acatcatagc caagcaaggc gataccttcg attctttggc cgaagagttc      720
gacatgagag cctccaaact ggccaaatac aacgatgctc ccgtggattt cccgatcgaa      780
aagggcgatg tgatctatct ggagaaaaag cagcatgct ccatctccaa acacacacag      840
cacgtagtgc gtgtgggcga ttcgatgcac agtatctccc aacgctatgg catccggatg      900
aagaacctct acaagctcaa cgacaaggat ggcgaatata taccccaaga gggcgatata      960
ctgcgcttgc gc                                         972

```

(2) INFORMATION FOR SEQ ID NO:10

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 201 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:  
(A) ORGANISM: *Porphyromonas gingivalis*
- (ix) FEATURE:  
(A) NAME/KEY: misc feature  
(B) LOCATION 1...201
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10

```

Glu Leu Ser Lys Cys Tyr Met Asp Lys Val Ser Tyr Ala Leu Gly Leu
1           5           10           15
Ser Ile Gly Asn Asn Phe Lys Ser Ser Gly Ile Asp Ser Val Val Met
          20          25          30
Asp Asp Phe Met Gln Gly Leu Ser Asp Val Leu Glu Glu Lys Ala Pro
          35          40          45
Gln Leu Ser Tyr Asp Glu Ala Lys Arg Glu Ile Glu Ala Tyr Phe Met
50          55          60

```

```

Asp Leu Gln Gln Lys Ala Val Lys Leu Asn Lys Glu Ala Gly Glu Glu
65      70      75      80
Phe Leu Lys Ile Asn Ala His Lys Glu Gly Val Thr Thr Leu Pro Ser
      85      90      95
Gly Leu Gln Tyr Glu Val Ile Lys Met Gly Glu Gly Pro Lys Pro Thr
      100     105     110
Leu Ser Asp Thr Val Thr Cys His Tyr His Gly Thr Leu Ile Asn Gly
      115     120     125
Ile Val Phe Asp Ser Ser Met Asp Arg Gly Glu Pro Ala Ser Phe Pro
      130     135     140
Leu Arg Gly Val Ile Ala Gly Trp Thr Glu Ile Leu Gln Leu Met Pro
      145     150     155     160
Val Gly Ser Lys Trp Lys Val Thr Ile Pro Ser Asp Leu Ala Tyr Gly
      165     170     175
Asp Arg Gly Ala Gly Glu His Ile Lys Pro Gly Ser Thr Leu Ile Phe
      180     185     190
Ile Ile Glu Leu Leu Ser Ile Asn Lys
      195     200

```

## (2) INFORMATION FOR SEQ ID NO:11

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 373 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...373

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11

```

Gly Ser Cys Gly Ser Ser Pro Ala Ile Gly Arg Ser Gly Cys Ser Arg
1      5      10      15
Cys Asp Arg Arg Ile Gly His Gly Leu Ser Thr Ser Ala Phe Arg Thr
      20      25      30
Tyr Ala Gly Leu Val Cys His Ser Arg Leu Arg Asp Gln Ser Ser Cys
      35      40      45
Arg Val Cys Gly Met Arg Pro Ile Gly Gln Leu Phe Phe Pro Ser Cys
      50      55      60
Pro Arg Arg Ser Ala Ser Asp Ala Gly Arg Thr Glu Arg Ile Gln Ser
      65      70      75      80
Pro Leu Gln Asn Leu Leu Gln Glu Met Gln Phe Ser Pro Thr Asn Arg
      85      90      95
Arg Asn Pro Phe Asp Asn Met Asn Ser Arg His Leu Thr Ile Thr Ile
      100     105     110
Ile Ala Gly Leu Ser Leu Phe Val Leu Thr Leu Gly Gly Cys Ser Val
      115     120     125
Ala Gln Gln Asp Thr Gln Trp Thr Leu Gly Gly Lys Leu Phe Thr Ser
      130     135     140
Ala Trp Ile Gln Arg Ser Ala Glu Tyr Gln Ala Leu Cys Ile Gln Ala
      145     150     155     160
Tyr Asn Ile Ala Thr Glu Arg Val Asp Ala Leu Pro Ala Glu Arg Lys
      165     170     175
Gln Gly Asp Arg Pro Tyr Ala Ile Val Thr Asp Ile Asp Glu Thr Ile
      180     185     190

```

[illegible]

(2) INFORMATION FOR SEQ ID NO:12

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 311 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Porphyromonas gingivalis

(ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...311

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

Pro 1	Tyr	Asn	Val	Gln 5	Ser	Ile	Ser	Asn	Lys 10	Thr	Ile	Lys	Lys	Gln 15	Met
Glu	Asn	Leu	Lys 20	Asn	Ile	Gln	Pro	Arg 25	Glu	Asp	Phe	Asn	Trp 30	Glu	Glu
Phe	Glu	Ala 35	Gly	Gly	Val	His	Ala 40	Ala	Val	Ser	Arg	Gln 45	Glu	Gln	Glu
Ala	Ala	Tyr	Asp	Lys	Thr	Leu	Asn 55	Thr	Ile	Lys	Glu	Lys	Glu	Val	Val
Met 65	Gly	Arg	Val	Thr	Ala 70	Ile	Asn	Lys	Arg	Glu	Val	Val	Ile	Asn 80	Val
Gly	Tyr	Lys	Ser	Glu 85	Gly	Val	Val	Pro	Ala 90	Thr	Glu	Phe	Arg	Tyr 95	Asn
Pro	Glu	Leu	Lys 100	Val	Gly	Asp	Glu	Val	Glu 105	Val	Tyr	Ile	Glu	Asn 110	Gln
Glu	Asp	Lys 115	Lys	Gly	Gln	Leu	Val 120	Leu	Ser	His	Arg	Lys 125	Ala	Arg	Ala
Ala	Arg 130	Ser	Trp	Glu	Arg	Val 135	Asn	Glu	Ala	Leu	Glu 140	Lys	Asp	Glu	Ile

```

Val Lys Gly Tyr Val Lys Cys Arg Thr Lys Gly Gly Met Ile Val Asp
145          150          155          160
Val Phe Gly Ile Glu Ala Phe Leu Pro Gly Ser Gln Ile Asp Val Arg
          165          170          175
Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr Met Glu Phe Lys
          180          185          190
Ile Val Lys Ile Asn Gln Glu Tyr Lys Asn Val Val Val Ser His Lys
          195          200          205
Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys Glu Ile Ile Gly
          210          215          220
Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val Lys Asn Ile Thr
225          230          235          240
Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp Gly Leu Ile His
          245          250          255
Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro Glu Glu Ile Val
          260          265          270
Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp Phe Asp Glu Asp
          275          280          285
Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met Pro Ser Ser Leu
290          295          300
Gly Cys Ser Arg Gln Arg Ala
305          310

```

## (2) INFORMATION FOR SEQ ID NO:13

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 985 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...985

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13

```

Val Thr Asp Lys Met Gln Asn Lys Gly Phe Val Ile Val Ile Thr Ser
1          5          10          15
Ala Leu Ala Ile Ile Cys Ala Phe Tyr Leu Ser Phe Ser Phe Val Thr
          20          25          30
Asn Arg Tyr Glu Lys Lys Ala Lys Ala Met Gly Asp Val Ala Gly Met
          35          40          45
Ala Tyr Leu Asp Ser Met Ser Asn Glu Lys Val Trp Phe Gly Tyr Thr
          50          55          60
Leu Lys Glu Ala Gln Ala Gln Gln Ile Gly Leu Gly Leu Asp Leu Lys
65          70          75          80
Gly Gly Met Asn Val Ile Leu Lys Leu Asn Ala Ser Asp Leu Leu Arg
          85          90          95
Asn Leu Ser Asn Lys Ser Leu Asp Pro Asn Phe Asn Lys Ala Leu Glu
          100          105          110
Asn Ala Ala Lys Ser Thr Glu Gln Ser Asp Phe Ile Asp Ile Phe Val
          115          120          125
Lys Glu Tyr Arg Lys Leu Asp Pro Asn Gly Arg Leu Ala Val Ile Phe
130          135          140
Gly Ser Gly Asp Leu Arg Asp Gln Ile Thr Ala Lys Ser Thr Asp Ala
145          150          155          160

```

Asp	Val	Val	Arg	Leu	Leu	Lys	Glu	Lys	Tyr	Asn	Ser	Ala	Val	Glu	Ala	
				165					170					175		
Ser	Phe	Asn	Val	Leu	Arg	Ala	Arg	Ile	Asp	Ala	Phe	Gly	Val	Val	Ala	
			180					185					190			
Pro	Asn	Leu	Gln	Arg	Leu	Glu	Gly	Gln	Gly	Arg	Ile	Leu	Val	Glu	Leu	
		195					200					205				
Pro	Gly	Val	Lys	Asp	Pro	Glu	Arg	Val	Arg	Thr	Leu	Leu	Gln	Arg	Ser	
	210					215					220					
Ala	Asn	Leu	Gln	Phe	Trp	Arg	Thr	Tyr	Lys	Phe	Glu	Glu	Val	Ser	Gly	
225					230					235					240	
Asp	Leu	Ile	Ala	Ala	Asn	Asp	Arg	Leu	Ser	Glu	Leu	Ala	Met	Asn	Asn	
				245					250					255		
Thr	Asp	Ala	Thr	Pro	Glu	Thr	Glu	Pro	Ala	Thr	Thr	Asp	Ser	Val	Ala	
			260					265					270			
Ala	Thr	Ala	Asp	Ser	Ala	Ala	Val	Gln	Ala	Val	Ala	Asp	Ser	Ala	Thr	
		275					280					285				
Val	Ala	Gln	Lys	Glu	Ala	Lys	Asp	Ala	Thr	Arg	Lys	Asp	Ala	Leu	Phe	
	290					295					300					
Ser	Leu	Leu	Thr	Pro	Val	Asn	Arg	Gly	Gly	Ala	Val	Val	Gly	Val	Ala	
305					310					315					320	
Arg	Arg	Ala	Asn	Met	Ala	Gln	Ile	Ser	Glu	Met	Leu	Gln	Gln	Ala	His	
			325						330					335		
Asp	Leu	Lys	Val	Thr	Arg	Glu	Asp	Val	Leu	Phe	Leu	Trp	Gly	Ala	Lys	
			340					345					350			
Ala	Ile	Glu	Asp	Pro	Glu	Thr	Lys	Lys	Glu	Thr	Asp	Leu	Tyr	Glu	Leu	
		355					360					365				
Tyr	Ala	Ile	Arg	Thr	Asn	Arg	Thr	Gly	Asp	Pro	Asp	Leu	Gly	Gly	Asp	
	370					375					380					
Val	Val	Thr	Ser	Ala	Lys	Ser	Asp	Ile	Gln	Asn	Asp	Phe	Gly	Arg	Ser	
385					390					395					400	
Glu	Pro	Ile	Val	Ser	Met	Thr	Met	Asn	Glu	Gly	Ala	Arg	Lys	Trp		
				405					410				415			
Ala	Arg	Ile	Thr	Lys	Asp	Asn	Val	Gly	Arg	Ala	Ile	Ala	Ile	Val	Leu	
			420					425					430			
Asp	Gly	Val	Val	Tyr	Ser	Ala	Pro	Asn	Val	Asn	Asp	Glu	Ile	Thr	Gly	
		435					440					445				
Gly	Arg	Ser	Gln	Ile	Ser	Gly	His	Phe	Thr	Val	Glu	Ala	Gly	Asp		
	450					455					460					
Leu	Ala	Asn	Val	Leu	Asn	Ser	Gly	Lys	Met	Asp	Ala	Thr	Val	Ser	Ile	
465					470					475					480	
Glu	Gln	Glu	Asn	Val	Ile	Gly	Pro	Thr	Leu	Gly	Ala	Glu	Ser	Ile	Lys	
			485						490					495		
Ala	Gly	Phe	Leu	Ser	Phe	Leu	Leu	Ala	Leu	Val	Ile	Leu	Met	Cys	Tyr	
			500					505					510			
Met	Cys	Leu	Ala	Tyr	Gly	Phe	Leu	Pro	Gly	Leu	Ile	Ala	Asn	Gly	Ala	
		515					520					525				
Leu	Ile	Val	Asn	Ser	Phe	Phe	Thr	Leu	Gly	Val	Leu	Ala	Ser	Phe	His	
	530					535					540					
Ala	Val	Leu	Thr	Leu	Ser	Gly	Ile	Ala	Gly	Leu	Val	Leu	Thr	Leu	Gly	
545					550					555					560	
Met	Ala	Val	Asp	Ala	Asn	Val	Leu	Ile	Phe	Glu	Arg	Ile	Lys	Glu	Glu	
				565					570					575		
Leu	Arg	Ala	Gly	Lys	Thr	Pro	Ile	Arg	Ala	Val	Thr	Asp	Gly	Tyr	Gly	
			580					585					590			
Asn	Ala	Phe	Ser	Ala	Ile	Phe	Asp	Ser	Asn	Val	Thr	Thr	Ile	Ile	Thr	
		595					600						605			
Gly	Ile	Ile	Leu	Phe	Leu	Tyr	Gly	Thr	Gly	Pro	Ile	Arg	Gly	Phe	Ala	
	610					615					620					
Thr	Thr	Leu	Ile	Ile	Gly	Leu	Ile	Ala	Ser	Phe	Ile	Thr	Ala	Val	Phe	
625					630					635					640	
Leu	Thr	Arg	Ile	Val	Phe	Glu	Lys	Leu	Ala	Lys	Lys	Gly	Arg	Leu	Asp	
				645					650					655		
Lys	Ile	Thr	Phe	Thr	Thr	Ser	Ile	Thr	Arg	Asn	Leu	Leu	Val	Asn	Pro	
			660					665					670			

```

Ser Tyr Asn Ile Leu Gly Lys Arg Lys Thr Gly Phe Ile Ile Pro Val
675                               680                               685
Ile Ile Ile Val Leu Gly Leu Ile Ala Ser Phe Thr Ile Gly Leu Asn
690                               695                               700
Arg Gly Ile Glu Phe Ser Gly Gly Arg Asn Tyr Val Val Lys Phe Asp
705                               710                               715                               720
Gln Pro Val Ser Ser Glu Ala Val Arg Ser Ala Leu Ser Ser Pro Leu
725                               730                               735
Gln Glu Lys Val Leu Val Thr Ser Ile Gly Thr Glu Gly Thr Glu Val
740                               745                               750
Arg Ile Ser Thr Asn Tyr Lys Ile Gln Glu Glu Ser Glu Glu Thr Glu
755                               760                               765
Ala Glu Ile Thr Asp Lys Leu Tyr Gln Ser Leu Lys Gly Phe Tyr Thr
770                               775                               780
Gln Gln Pro Thr Ala Asp Gln Phe Leu Asp Asn Ile Ile Ser Ser Gln
785                               790                               795                               800
Lys Val Ser Pro Ser Met Ser Ser Asp Ile Thr Arg Gly Ala Ile Trp
805                               810                               815
Ala Val Leu Leu Ser Met Ile Phe Met Ala Ile Tyr Ile Leu Ile Arg
820                               825                               830
Phe Arg Asp Ile Ser Phe Ser Ala Gly Val Phe Val Ser Val Ala Ala
835                               840                               845
Thr Thr Phe Cys Ile Ile Ala Leu Tyr Ala Leu Leu Trp Lys Ile Leu
850                               855                               860
Pro Phe Thr Met Glu Ile Asp Gln Asn Phe Ile Ala Ala Ile Leu Ala
865                               870                               875                               880
Ile Ile Gly Tyr Ser Leu Asn Asp Thr Val Val Val Phe Asp Arg Ile
885                               890                               895
Arg Glu Thr Met Lys Leu Tyr Pro Asn Arg Asp Arg Tyr Gln Val Ile
900                               905                               910
Asn Asp Ala Leu Asn Ser Thr Leu Gly Arg Thr Leu Asn Thr Ser Leu
915                               920                               925
Thr Thr Phe Ile Val Met Leu Val Ile Phe Ile Phe Gly Gly Ala Thr
930                               935                               940
Met Arg Ser Phe Thr Phe Ser Ile Leu Leu Gly Ile Val Ile Gly Thr
945                               950                               955                               960
Tyr Ser Thr Leu Phe Val Ala Thr Pro Leu Ala Tyr Glu Ile Gln Lys
965                               970                               975
Arg Lys Leu Asn Lys Ala Ala Lys Lys
980                               985

```

## (2) INFORMATION FOR SEQ ID NO:14

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 106 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...106

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

```

Leu Ser Tyr Ser Gly Glu Ser Asp Ala Lys Glu Ser Asp Gln Asn Cys
1           5           10           15

```

Arg	Lys	Cys	Thr	Phe	Ile	Gly	Phe	Glu	Lys	Arg	Val	Asn	Thr	Met	Arg
			20					25					30		
Leu	Ile	Lys	Ala	Phe	Leu	Val	Gln	Leu	Leu	Leu	Leu	Pro	Ile	Phe	Phe
		35					40					45			
Tyr	Lys	Arg	Phe	Ile	Ser	Pro	Leu	Thr	Pro	Pro	Ser	Cys	Arg	Phe	Thr
	50					55					60				
Pro	Ser	Cys	Ser	Ser	Tyr	Ala	Ile	Glu	Ala	Leu	Arg	Lys	Tyr	Gly	Pro
65					70					75					80
Gly	Lys	Gly	Leu	Leu	Leu	Ser	Ile	Lys	Arg	Ile	Leu	Arg	Cys	His	Pro
			85						90					95	
Trp	Gly	Gly	Ser	Gly	Tyr	Asp	Pro	Val	Pro						
			100					105							

## (2) INFORMATION FOR SEQ ID NO:15

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 456 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...456

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

Arg	Gly	Gly	Gln	Ile	Arg	Arg	His	His	Thr	Asp	Ser	Ser	Arg	Gly	Ser
1				5					10					15	
Asp	Ser	Lys	Ala	Cys	Arg	Phe	Ser	Phe	His	Ala	Glu	Thr	Ile	Gly	Phe
			20					25					30		
Ser	Asn	His	Gln	Arg	Ala	Lys	Met	Asn	Tyr	Leu	Tyr	Ile	Leu	Ile	Thr
		35					40					45			
Leu	Leu	Leu	Ser	Gly	Phe	Phe	Ser	Gly	Ala	Glu	Ile	Ala	Phe	Leu	Ser
	50					55				60					
Ser	Asp	Lys	Leu	Arg	Leu	Glu	Leu	Asp	Arg	Asn	Arg	Gly	Asp	Leu	Thr
65					70					75					80
Gly	Arg	Ala	Leu	Asn	Leu	Leu	Tyr	Arg	His	Pro	Asp	Gln	Leu	Val	Thr
			85						90					95	
Thr	Leu	Leu	Val	Gly	Asn	Asn	Ile	Val	Leu	Val	Val	Tyr	Gly	Leu	Leu
			100					105					110		
Met	Ala	Gly	Leu	Leu	Ala	Ala	Pro	Leu	Ala	Gln	Trp	Ile	Asp	Asn	Asp
	115						120					125			
Ala	Met	Ile	Val	Val	Leu	Gln	Ser	Val	Leu	Ser	Thr	Ile	Ile	Ile	Leu
	130					135					140				
Phe	Thr	Gly	Glu	Phe	Leu	Pro	Lys	Ala	Ile	Phe	Lys	Thr	Asn	Ala	Asn
145					150					155					160
Met	Met	Met	Arg	Val	Phe	Ala	Leu	Pro	Ile	Val	Ala	Ile	Tyr	Tyr	Leu
			165						170					175	
Leu	Tyr	Pro	Leu	Ser	Lys	Leu	Phe	Thr	Gly	Leu	Ser	Arg	Ser	Phe	Ile
			180					185					190		
Arg	Leu	Val	Asp	Lys	Asn	Tyr	Val	Pro	Thr	Thr	Val	Gly	Leu	Gly	Arg
		195					200					205			
Val	Asp	Leu	Asp	His	Tyr	Leu	Ala	Glu	Asn	Met	Ser	Gly	Glu	Asn	Glu
	210					215					220				
Gln	Asn	Asp	Leu	Thr	Thr	Glu	Val	Lys	Ile	Ile	Gln	Asn	Ala	Leu	Asp
225					230					235					240



```

Phe Ser Gly Ile Gln Val Arg Asp Cys Met Ile Pro Arg Asn Glu Met
      245                250                255
Ile Ala Cys Glu Leu Gln Thr Asp Ile Glu Val Leu Lys Thr Thr Phe
      260                265                270
Ile Asp Thr Gly Leu Ser Lys Ile Ile Ile Tyr Arg Gln Asn Ile Asp
      275                280                285
Asp Val Val Gly Tyr Ile His Ser Ser Glu Met Phe Arg Gly Gln Asp
      290                295                300
Trp Gln Lys Arg Ile Asn Thr Thr Val Phe Val Pro Glu Ser Met Tyr
      305                310                315
Ala Asn Lys Leu Met Arg Leu Leu Met Gln Arg Lys Lys Ser Ile Ala
      325                330                335
Ile Val Ile Asp Glu Leu Gly Gly Thr Ala Gly Met Val Thr Leu Glu
      340                345                350
Asp Leu Val Glu Glu Ile Phe Gly Asp Ile Glu Asp Glu His Asp Thr
      355                360                365
Arg Lys Ile Ile Ala Lys Gln Leu Gly Pro His Thr Tyr Leu Val Ser
      370                375                380
Gly Arg Met Glu Ile Asp Asp Val Asn Glu Arg Phe Gly Leu Ser Leu
      385                390                395
Pro Glu Ser Asp Asp Tyr Leu Thr Val Ala Gly Phe Ile Leu Asn Ser
      405                410                415
His Gln Asn Ile Pro Gln Ala Asn Glu Val Val Glu Ile Ala Pro Tyr
      420                425                430
Thr Phe Thr Ile Leu Arg Ser Ser Thr Lys Ile Glu Leu Val Lys
      435                440                445
Met Ser Ile Asp Asp Gln Ser Asn
      450                455

```

## (2) INFORMATION FOR SEQ ID NO:16

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 679 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...679

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

```

Pro Ile Pro Phe Leu Leu Asn Met Tyr Ser Gly His His Lys Ile His
1      5      10      15
Tyr Pro Phe Leu Ile Leu Leu Val Cys Leu Ala Phe Ala Ala Cys Lys
      20      25      30
Ser Val Lys Leu Lys Asp Ala Glu Lys Ala His Asp Arg Gln Glu Tyr
      35      40      45
Thr Lys Ala Ala Asp Met Tyr Asn Thr Leu Tyr Arg Arg Thr Arg Arg
      50      55      60
Lys Gln Val Glu Met Lys Ala Tyr Thr Ala Phe Arg Ser Gly Glu Asn
      65      70      75      80
Tyr Arg Ala Ala Gly Arg Gln Ala Lys Ala Leu Arg Gly Tyr Leu Asn
      85      90      95
Ala Arg Arg Tyr Gly Tyr Pro Asp Ser Val Val Leu Leu Arg Leu Ala
      100     105     110

```

Gln	Thr	Tyr	Gln	Gln	Gly	Gly	Asn	Tyr	Lys	Glu	Ala	Glu	Val	Leu	Phe		
		115					120					125					
Arg	Gly	Tyr	Leu	Glu	Ala	Tyr	Pro	Lys	Ser	Tyr	Phe	Ala	Ala	Ile	Gly		
	130					135					140						
Leu	Glu	Gly	Cys	Leu	Phe	Ala	Arg	Gln	Gln	Lys	Glu	Tyr	Pro	Thr	Arg		
145					150					155					160		
Tyr	Arg	Ile	Arg	Arg	Ala	Ala	Glu	Trp	Asn	Ser	Ala	Arg	Gly	Asp	Phe		
				165					170					175			
Gly	Pro	Ala	Tyr	Ala	Pro	Asp	Ala	Ser	Ala	Leu	Tyr	Phe	Thr	Ser	Ser		
		180						185					190				
Arg	Ser	Lys	Asp	Asp	Gly	Leu	Asp	Asn	Ser	Ser	Ile	Thr	Gly	Leu	Lys		
		195					200					205					
Pro	Asn	Asp	Ile	Tyr	Ile	Ile	Lys	Arg	Asp	Ala	Gln	Gly	Arg	Trp	Gly		
	210					215					220						
Arg	Pro	Asp	Ser	Val	Ser	Gly	Gly	Ile	Asn	Thr	Pro	Trp	Asp	Glu	Gly		
225					230					235					240		
Val	Pro	Thr	Ile	Thr	Pro	Asp	Gly	Ser	Thr	Ile	Tyr	Tyr	Thr	Leu	Ala		
				245					250					255			
Gln	Gln	Gly	Ala	Asp	Tyr	Asp	Arg	Thr	Val	Gln	Ile	Tyr	Ser	Ala	Ala		
			260					265					270				
Arg	Ser	Gly	Glu	Gly	Gly	Trp	Ser	Asn	Gly	Ser	Leu	Val	Asp	Ile	Met		
		275					280					285					
Arg	Asp	Ser	Leu	Arg	Met	Ala	Ala	His	Pro	Ser	Met	Ser	Ala	Ser	Gly		
	290					295					300						
Asp	Tyr	Leu	Tyr	Phe	Val	Ser	Asn	Ile	Gly	Gly	Ser	Tyr	Gly	Gly	Lys		
305					310					315					320		
Asp	Ile	Tyr	Arg	Val	Lys	Val	Ser	Asp	Arg	Ser	Tyr	Gly	Ser	Pro	Glu		
				325					330					335			
Asn	Leu	Gly	Pro	Asp	Ile	Asn	Thr	Pro	Gly	Asp	Glu	Met	Phe	Pro	Phe		
			340					345					350				
Ile	Asp	Gly	Asp	Ser	Thr	Leu	Phe	Ala	Ser	Asp	Gly	His	Ala	Gly			
		355					360					365					
Leu	Gly	Gly	Leu	Asp	Ile	Phe	Lys	Ala	Thr	Leu	Asp	Ser	Thr	Gly	Gln		
	370					375					380						
Trp	His	Val	Val	Asn	Met	Gly	Gln	Pro	Val	Asn	Ser	Ser	Ala	Asp	Asp		
385					390					395				400			
Phe	Gly	Leu	Ala	Val	Glu	Pro	Lys	Gly	Lys	Asn	Lys	Glu	Glu	Ala	Leu		
				405					410					415			
Pro	Asp	Asn	Gly	Val	Lys	Gly	Val	Phe	Cys	Ser	Asn	Arg	Gly	Asp	Ala		
			420					425					430				
Arg	Gly	Trp	Pro	His	Leu	Phe	His	Phe	Glu	Leu	Pro	Ala	Ile	Tyr	Thr		
		435					440					445					
Glu	Ile	Gln	Gly	Tyr	Val	Met	Asp	Arg	Glu	Glu	Asn	Pro	Ile	Ala	Gly		
	450					455					460						
Ala	Thr	Val	Arg	Ile	Val	Gly	Glu	Arg	Gly	Pro	Val	Gly	Gln	Gly	Phe		
465					470					475					480		
Val	Thr	Thr	Arg	Asp	Asp	Gly	Ser	Tyr	Lys	Met	Ser	Val	Gln	Gly	Asp		
				485					490					495			
Thr	Arg	Tyr	Val	Met	Leu	Ala	Gly	Ala	Ser	Gly	Tyr	Leu	Asn	Gln	Tyr		
			500					505					510				
Val	Glu	Leu	Lys	Thr	Asp	Thr	Ala	Lys	Gln	Ser	Glu	Thr	Tyr	Tyr	Val		
		515					520					525					
Asp	Phe	Phe	Leu	Ala	Ser	Arg	Glu	Lys	Ala	Glu	Gly	Leu	Gln	Asn	Ile		
	530					535					540						
Phe	Tyr	Asp	Phe	Asp	Lys	Ala	Thr	Leu	Arg	Pro	Glu	Ser	Met	Lys	Ser		
545					550					555				560			
Leu	Asp	Glu	Leu	Ile	Arg	Ile	Leu	Thr	Asp	Asn	Pro	Asp	Ile	Arg	Ile		
				565					570					575			
Glu	Leu	Gly	Ser	His	Ala	Asp	Arg	Lys	Gly	Pro	Asp	Ala	Tyr	Asn	Leu		
			580					585					590				
Gly	Leu	Ser	Asp	Arg	Arg	Ala	Lys	Ser	Val	Val	Asp	Tyr	Leu	Thr	Ser		
		595					600					605					
Arg	Gly	Ile	Ala	Ala	Asp	Arg	Leu	Thr	Trp	Lys	Gly	Tyr	Gly	Lys	Ser		
	610					615					620						

```

Val Pro Lys Thr Val Thr Ala Lys Ile Ala Glu Arg His Asp Phe Leu
625                               630           635           640
Lys Glu Gly Asp Val Leu Thr Glu Glu Phe Val Ala Pro Leu Thr Glu
                               645           650           655
Glu Gln Gln Ser Val Cys Asp Gln Leu Asn Arg Arg Thr Glu Phe Arg
660                               665           670
Val Ile Glu Glu Glu Leu Arg
675

```

## (2) INFORMATION FOR SEQ ID NO:17

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 337 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature
- (B) LOCATION 1...337

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

```

Pro Gln Pro Val Gly Leu Lys Glu Ile Thr Ile Lys Pro Met Cys Leu
1                               5           10           15
Glu Pro Ile Ile Ala Pro Ile Ser Ser Glu Leu Leu Glu Gln Glu Leu
20                               25           30
Thr Ala Asp Arg Phe Leu Arg Met Thr Asn Lys Ala Gly Asn Glu Ile
35                               40           45
Tyr Val Phe Thr Ala Glu Glu Ala Pro His Cys Met Lys Glu Val Gly
50                               55           60
Arg Leu Arg Glu Glu Ala Phe Arg His Tyr Gly Gly Gly Thr Gly Lys
65                               70           75           80
Ala Ile Asp Ile Asp Glu Phe Asp Thr Met Pro Gly Ser Tyr Lys Gln
85                               90           95
Leu Ile Val Trp Asp Pro Gln Asn Lys Ala Ile Leu Gly Gly Tyr Arg
100                              105          110
Phe Ile Tyr Gly Arg Asp Val Ala Phe Asp Thr Asp Gly Lys Pro Leu
115                              120          125
Leu Ala Thr Ala Glu Met Phe Arg Phe Ser Asp Ala Phe Leu His Asp
130                              135          140
Tyr Leu Pro Tyr Thr Val Glu Leu Gly Arg Ser Phe Val Ser Leu Gln
145                              150          155          160
Tyr Gln Ser Thr Arg Met Gly Thr Lys Ala Ile Phe Val Leu Asp Asn
165                              170          175
Leu Trp Asp Gly Ile Gly Ala Leu Thr Val Val Asn Pro Glu Ala Leu
180                              185          190
Tyr Phe Tyr Gly Lys Val Thr Met Tyr Lys Asp Tyr Asp Arg Arg Ala
195                              200          205
Arg Asn Leu Ile Leu Tyr Phe Leu Arg Lys His Phe Ser Asp Pro Glu
210                              215          220
Gly Leu Val Lys Pro Ile His Pro Leu Pro Ile Glu Ile Ser Ala Glu
225                              230          235          240
Asp Glu Ala Leu Phe Ser Ser Ser Asp Phe Asp Thr Asn Tyr Lys Thr
245                              250          255
Leu Asn Ile Glu Val Arg Lys Leu Gly Ile Asn Ile Pro Pro Leu Val
260                              265          270

```

Ser Ala Tyr Ile Ala Leu Ser Pro Glu Met Arg Val Phe Gly Thr Ala  
 275 280 285  
 Val Asn Glu Ser Phe Gly Glu Val Glu Glu Thr Gly Ile Phe Ile Ala  
 290 295 300  
 Val Gly Lys Ile Leu Glu Glu Lys Lys Gln Arg His Ile Glu Ser Phe  
 305 310 315 320  
 Ile Leu Ser Arg Asn Glu Lys Lys Gly Leu Asp Ser Ser Asn Gly Arg  
 325 330 335  
 Ser

## (2) INFORMATION FOR SEQ ID NO:18

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 324 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...324

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

Leu Thr Leu Phe Cys Cys Asn Tyr Phe Lys Gln Met Arg Ala Asn Ile  
 1 5 10 15  
 Trp Gln Ile Leu Ser Val Ser Val Leu Phe Phe Phe Gly Thr Ala Ile  
 20 25 30  
 Gly Gln Ala Gln Ser Arg Asn Arg Thr Tyr Glu Ala Tyr Val Lys Gln  
 35 40 45  
 Tyr Ala Asp Glu Ala Ile Arg Gln Met Ser Arg Tyr Asn Ile Pro Ala  
 50 55 60  
 Ser Ile Thr Ile Ala Gln Ala Leu Val Glu Thr Gly Ala Gly Ala Ser  
 65 70 75 80  
 Thr Leu Ala Ser Val His Asn Asn His Phe Gly Ile Lys Cys His Lys  
 85 90 95  
 Ser Trp Thr Gly Lys Arg Thr Tyr Arg Thr Asp Asp Ala Pro Asn Glu  
 100 105 110  
 Cys Phe Arg Ser Tyr Ser Ala Ala Arg Glu Ser Tyr Glu Asp His Ser  
 115 120 125  
 Arg Phe Leu Leu Gln Pro Arg Tyr Arg Pro Leu Phe Lys Leu Asp Arg  
 130 135 140  
 Glu Asp Tyr Arg Gly Trp Ala Thr Gly Leu Gln Arg Cys Gly Tyr Ala  
 145 150 155 160  
 Thr Asn Arg Gly Tyr Ala Asn Leu Leu Ile Lys Met Val Glu Leu Tyr  
 165 170 175  
 Glu Leu Tyr Ala Leu Asp Arg Glu Lys Tyr Pro Ser Trp Phe His Lys  
 180 185 190  
 Ser Tyr Pro Gly Ser Asn Lys Lys Ser His Gln Thr Thr Lys Gln Lys  
 195 200 205  
 Gln Ser Gly Leu Lys His Glu Ala Tyr Phe Ser Tyr Gly Leu Leu Tyr  
 210 215 220  
 Ile Ile Ala Lys Gln Gly Asp Thr Phe Asp Ser Leu Ala Glu Glu Phe  
 225 230 235 240  
 Asp Met Arg Ala Ser Lys Leu Ala Lys Tyr Asn Asp Ala Pro Val Asp  
 245 250 255  
 Phe Pro Ile Glu Lys Gly Asp Val Ile Leu Glu Lys Lys His Ala

```

          260          265          270
Cys Ser Ile Ser Lys His Thr Gln His Val Val Arg Val Gly Asp Ser
          275          280          285
Met His Ser Ile Ser Gln Arg Tyr Gly Ile Arg Met Lys Asn Leu Tyr
          290          295          300
Lys Leu Asn Asp Lys Asp Gly Glu Tyr Ile Pro Gln Glu Gly Asp Ile
305          310          315          320
Leu Arg Leu Arg

```

## (2) INFORMATION FOR SEQ ID NO:19

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 195 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...195

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

```

Met Asp Lys Val Ser Tyr Ala Leu Gly Leu Ser Ile Gly Asn Asn Phe
1          5          10          15
Lys Ser Ser Gly Ile Asp Ser Val Val Met Asp Asp Phe Met Gln Gly
          20          25          30
Leu Ser Asp Val Leu Glu Glu Lys Ala Pro Gln Leu Ser Tyr Asp Glu
          35          40          45
Ala Lys Arg Glu Ile Glu Ala Tyr Phe Met Asp Leu Gln Gln Lys Ala
50          55          60
Val Lys Leu Asn Lys Glu Ala Gly Glu Glu Phe Leu Lys Ile Asn Ala
65          70          75          80
His Lys Glu Gly Val Thr Thr Leu Pro Ser Gly Leu Gln Tyr Glu Val
          85          90          95
Ile Lys Met Gly Glu Gly Pro Lys Pro Thr Leu Ser Asp Thr Val Thr
100          105          110
Cys His Tyr His Gly Thr Leu Ile Asn Gly Ile Val Phe Asp Ser Ser
115          120          125
Met Asp Arg Gly Glu Pro Ala Ser Phe Pro Leu Arg Gly Val Ile Ala
130          135          140
Gly Trp Thr Glu Ile Leu Gln Leu Met Pro Val Gly Ser Lys Trp Lys
145          150          155          160
Val Thr Ile Pro Ser Asp Leu Ala Tyr Gly Asp Arg Gly Ala Gly Glu
          165          170          175
His Ile Lys Pro Gly Ser Thr Leu Ile Phe Ile Ile Glu Leu Leu Ser
180          185          190
Ile Asn Lys
195

```

## (2) INFORMATION FOR SEQ ID NO:20

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 271 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Porphyromonas gingivalis
- (ix) FEATURE:
  - (A) NAME/KEY: misc feature
  - (B) LOCATION 1...271
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

```

Met Asn Ser Arg His Leu Thr Ile Thr Ile Ile Ala Gly Leu Ser Leu
1      5      10      15
Phe Val Leu Thr Leu Gly Gly Cys Ser Val Ala Gln Gln Asp Thr Gln
20      25      30
Trp Thr Leu Gly Gly Lys Leu Phe Thr Ser Ala Trp Ile Gln Arg Ser
35      40      45
Ala Glu Tyr Gln Ala Leu Cys Ile Gln Ala Tyr Asn Ile Ala Thr Glu
50      55      60
Arg Val Asp Ala Leu Pro Ala Glu Arg Lys Gln Gly Asp Arg Pro Tyr
65      70      75      80
Ala Ile Val Thr Asp Ile Asp Glu Thr Ile Leu Asp Asn Thr Pro Asn
85      90      95
Ser Val Tyr Gln Ala Leu Arg Gly Lys Asp Tyr Asp Glu Glu Thr Trp
100     105     110
Gly Lys Trp Cys Ala Gln Ala Asp Ala Asp Thr Leu Ala Gly Ala Leu
115     120     125
Ser Phe Phe Leu His Ala Ala Asn Lys Gly Ile Glu Val Phe Tyr Val
130     135     140
Thr Asn Arg Arg Asp Asn Leu Arg Glu Ala Thr Leu Gln Asn Leu Gln
145     150     155     160
Arg Tyr Gly Phe Pro Phe Ala Asp Glu Glu His Leu Leu Thr Thr His
165     170     175
Gly Pro Ser Asp Lys Glu Pro Arg Arg Leu Lys Ile Gln Glu Gln Tyr
180     185     190
Glu Ile Val Leu Leu Ile Gly Asp Asn Leu Gly Asp Phe His His Phe
195     200     205
Phe Asn Thr Lys Glu Glu Ser Gly Arg Lys Gln Ala Leu Gly Leu Thr
210     215     220
Ala Gly Glu Phe Gly Arg His Phe Ile Met Leu Pro Asn Pro Asn Tyr
225     230     235     240
Gly Ser Trp Glu Pro Ala Trp Tyr Gly Gly Lys Tyr Pro Pro Leu Pro
245     250     255
Glu Arg Asp Lys Ala Leu Lys Gln Leu His Ser Gln Asn Ser Arg
260     265     270

```

(2) INFORMATION FOR SEQ ID NO:21

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 296 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein
- (iii) HYPOTHETICAL: YES
- (vi) ORIGINAL SOURCE:
  - (A) ORGANISM: Porphyromonas gingivalis

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...296

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

```

Met Glu Asn Leu Lys Asn Ile Gln Pro Arg Glu Asp Phe Asn Trp Glu
1      5      10      15
Glu Phe Glu Ala Gly Gly Val His Ala Ala Val Ser Arg Gln Glu Gln
20      25      30
Glu Ala Ala Tyr Asp Lys Thr Leu Asn Thr Ile Lys Glu Lys Glu Val
35      40      45
Val Met Gly Arg Val Thr Ala Ile Asn Lys Arg Glu Val Val Ile Asn
50      55      60
Val Gly Tyr Lys Ser Glu Gly Val Val Pro Ala Thr Glu Phe Arg Tyr
65      70      75      80
Asn Pro Glu Leu Lys Val Gly Asp Glu Val Glu Val Tyr Ile Glu Asn
85      90      95
Gln Glu Asp Lys Lys Gly Gln Leu Val Leu Ser His Arg Lys Ala Arg
100     105     110
Ala Ala Arg Ser Trp Glu Arg Val Asn Glu Ala Leu Glu Lys Asp Glu
115     120     125
Ile Val Lys Gly Tyr Val Lys Cys Arg Thr Lys Gly Gly Met Ile Val
130     135     140
Asp Val Phe Gly Ile Glu Ala Phe Leu Pro Gly Ser Gln Ile Asp Val
145     150     155     160
Arg Pro Ile Arg Asp Tyr Asp Ala Phe Val Glu Lys Thr Met Glu Phe
165     170     175
Lys Ile Val Lys Ile Asn Gln Glu Tyr Lys Asn Val Val Val Ser His
180     185     190
Lys Val Leu Ile Glu Ala Glu Leu Glu Gln Gln Lys Lys Glu Ile Ile
195     200     205
Gly Lys Leu Glu Lys Gly Gln Val Leu Glu Gly Ile Val Lys Asn Ile
210     215     220
Thr Ser Tyr Gly Val Phe Ile Asp Leu Gly Gly Val Asp Gly Leu Ile
225     230     235     240
His Ile Thr Asp Leu Ser Trp Gly Arg Val Ala His Pro Glu Glu Ile
245     250     255
Val Gln Leu Asp Gln Lys Ile Asn Val Val Ile Leu Asp Phe Asp Glu
260     265     270
Asp Arg Lys Arg Ile Ala Leu Gly Leu Lys Gln Leu Met Pro Ser Ser
275     280     285
Leu Gly Cys Ser Arg Gln Arg Ala
290     295

```

## (2) INFORMATION FOR SEQ ID NO:22

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 981 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: Porphyromonas gingivalis

## (ix) FEATURE:

- (A) NAME/KEY: misc\_feature  
(B) LOCATION 1...981

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22

Met	Gln	Asn	Lys	Gly	Phe	Val	Ile	Val	Ile	Thr	Ser	Ala	Leu	Ala	Ile	1	5	10	15
Ile	Cys	Ala	Phe	Tyr	Leu	Ser	Phe	Ser	Phe	Val	Thr	Asn	Arg	Tyr	Glu	20	25	30	
Lys	Lys	Ala	Lys	Ala	Met	Gly	Asp	Val	Ala	Gly	Met	Ala	Tyr	Leu	Asp	35	40	45	
Ser	Met	Ser	Asn	Glu	Lys	Val	Trp	Phe	Gly	Tyr	Thr	Leu	Lys	Glu	Ala	50	55	60	
Gln	Ala	Gln	Gln	Ile	Gly	Leu	Gly	Leu	Asp	Leu	Lys	Gly	Gly	Met	Asn	65	70	75	80
Val	Ile	Leu	Lys	Leu	Asn	Ala	Ser	Asp	Leu	Leu	Arg	Asn	Leu	Ser	Asn	85	90	95	
Lys	Ser	Leu	Asp	Pro	Asn	Phe	Asn	Lys	Ala	Leu	Glu	Asn	Ala	Ala	Lys	100	105	110	
Ser	Thr	Glu	Gln	Ser	Asp	Phe	Ile	Asp	Ile	Phe	Val	Lys	Glu	Tyr	Arg	115	120	125	
Lys	Leu	Asp	Pro	Asn	Gly	Arg	Leu	Ala	Val	Ile	Phe	Gly	Ser	Gly	Asp	130	135	140	
Leu	Arg	Asp	Gln	Ile	Thr	Ala	Lys	Ser	Thr	Asp	Ala	Asp	Val	Val	Arg	145	150	155	160
Leu	Leu	Lys	Glu	Lys	Tyr	Asn	Ser	Ala	Val	Glu	Ala	Ser	Phe	Asn	Val	165	170	175	
Leu	Arg	Ala	Arg	Ile	Asp	Ala	Phe	Gly	Val	Val	Ala	Pro	Asn	Leu	Gln	180	185	190	
Arg	Leu	Glu	Gly	Gln	Gly	Arg	Ile	Leu	Val	Glu	Leu	Pro	Gly	Val	Lys	195	200	205	
Asp	Pro	Glu	Arg	Val	Arg	Thr	Leu	Leu	Gln	Arg	Ser	Ala	Asn	Leu	Gln	210	215	220	
Phe	Trp	Arg	Thr	Tyr	Lys	Phe	Glu	Glu	Val	Ser	Gly	Asp	Leu	Ile	Ala	225	230	235	240
Ala	Asn	Asp	Arg	Leu	Ser	Glu	Leu	Ala	Met	Asn	Asn	Thr	Asp	Ala	Thr	245	250	255	
Pro	Glu	Thr	Glu	Pro	Ala	Thr	Thr	Asp	Ser	Val	Ala	Ala	Thr	Ala	Asp	260	265	270	
Ser	Ala	Ala	Val	Gln	Ala	Val	Ala	Asp	Ser	Ala	Thr	Val	Ala	Gln	Lys	275	280	285	
Glu	Ala	Lys	Asp	Ala	Thr	Arg	Lys	Asp	Ala	Leu	Phe	Ser	Leu	Leu	Thr	290	295	300	
Pro	Val	Asn	Arg	Gly	Gly	Ala	Val	Val	Gly	Val	Ala	Arg	Arg	Ala	Asn	305	310	315	320
Met	Ala	Gln	Ile	Ser	Glu	Met	Leu	Gln	Gln	Ala	His	Asp	Leu	Lys	Val	325	330	335	
Thr	Arg	Glu	Asp	Val	Leu	Phe	Leu	Trp	Gly	Ala	Lys	Ala	Ile	Glu	Asp	340	345	350	
Pro	Glu	Thr	Lys	Lys	Glu	Thr	Asp	Leu	Tyr	Glu	Leu	Tyr	Ala	Ile	Arg	355	360	365	
Thr	Asn	Arg	Thr	Gly	Asp	Pro	Asp	Leu	Gly	Gly	Asp	Val	Val	Thr	Ser	370	375	380	
Ala	Lys	Ser	Asp	Ile	Gln	Asn	Asp	Phe	Gly	Arg	Ser	Glu	Pro	Ile	Val	385	390	395	400
Ser	Met	Thr	Met	Asn	Glu	Glu	Gly	Ala	Arg	Lys	Trp	Ala	Arg	Ile	Thr	405	410	415	
Lys	Asp	Asn	Val	Gly	Arg	Ala	Ile	Ala	Ile	Val	Leu	Asp	Gly	Val	Val	420	425	430	
Tyr	Ser	Ala	Pro	Asn	Val	Asn	Asp	Glu	Ile	Thr	Gly	Gly	Arg	Ser	Gln	435	440	445	
Ile	Ser	Gly	His	Phe	Thr	Val	Glu	Glu	Ala	Gly	Asp	Leu	Ala	Asn	Val	450	455	460	
Leu	Asn	Ser	Gly	Lys	Met	Asp	Ala	Thr	Val	Ser	Ile	Glu	Gln	Glu	Asn	465	470	475	480
Val	Ile	Gly	Pro	Thr	Leu	Gly	Ala	Glu	Ser	Ile	Lys	Ala	Gly	Phe	Leu	485	490	495	





## (2) INFORMATION FOR SEQ ID NO:23

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 76 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...76

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23

```

Met Arg Leu Ile Lys Ala Phe Leu Val Gln Leu Leu Leu Leu Pro Ile
1          5          10          15
Phe Phe Tyr Lys Arg Phe Ile Ser Pro Leu Thr Pro Pro Ser Cys Arg
          20          25          30
Phe Thr Pro Ser Cys Ser Ser Tyr Ala Ile Glu Ala Leu Arg Lys Tyr
          35          40          45
Gly Pro Gly Lys Gly Leu Leu Leu Ser Ile Lys Arg Ile Leu Arg Cys
          50          55          60
His Pro Trp Gly Gly Ser Gly Tyr Asp Pro Val Pro
65          70          75

```

## (2) INFORMATION FOR SEQ ID NO:24

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 417 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

(A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

(A) NAME/KEY: misc feature

(B) LOCATION 1...417

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24

```

Met Asn Tyr Leu Tyr Ile Leu Ile Thr Leu Leu Leu Ser Gly Phe Phe
1          5          10          15
Ser Gly Ala Glu Ile Ala Phe Leu Ser Ser Asp Lys Leu Arg Leu Glu
          20          25          30
Leu Asp Arg Asn Arg Gly Asp Leu Thr Gly Arg Ala Leu Asn Leu Leu
          35          40          45
Tyr Arg His Pro Asp Gln Leu Val Thr Thr Leu Leu Val Gly Asn Asn
          50          55          60
Ile Val Leu Val Val Tyr Gly Leu Leu Met Ala Gly Leu Leu Ala Ala
65          70          75          80
Pro Leu Ala Gln Trp Ile Asp Asn Asp Ala Met Ile Val Val Leu Gln
          85          90          95

```

```

Ser Val Leu Ser Thr Ile Ile Ile Leu Phe Thr Gly Glu Phe Leu Pro
      100      105      110
Lys Ala Ile Phe Lys Thr Asn Ala Asn Met Met Met Arg Val Phe Ala
      115      120      125
Leu Pro Ile Val Ala Ile Tyr Tyr Leu Leu Tyr Pro Leu Ser Lys Leu
      130      135      140
Phe Thr Gly Leu Ser Arg Ser Phe Ile Arg Leu Val Asp Lys Asn Tyr
      145      150      155      160
Val Pro Thr Thr Val Gly Leu Gly Arg Val Asp Leu Asp His Tyr Leu
      165      170      175
Ala Glu Asn Met Ser Gly Glu Asn Glu Gln Asn Asp Leu Thr Thr Glu
      180      185      190
Val Lys Ile Ile Gln Asn Ala Leu Asp Phe Ser Gly Ile Gln Val Arg
      195      200      205
Asp Cys Met Ile Pro Arg Asn Glu Met Ile Ala Cys Glu Leu Gln Thr
      210      215      220
Asp Ile Glu Val Leu Lys Thr Thr Phe Ile Asp Thr Gly Leu Ser Lys
      225      230      235      240
Ile Ile Ile Tyr Arg Gln Asn Ile Asp Asp Val Val Gly Tyr Ile His
      245      250      255
Ser Ser Glu Met Phe Arg Gly Gln Asp Trp Gln Lys Arg Ile Asn Thr
      260      265      270
Thr Val Phe Val Pro Glu Ser Met Tyr Ala Asn Lys Leu Met Arg Leu
      275      280      285
Leu Met Gln Arg Lys Lys Ser Ile Ala Ile Val Ile Asp Glu Leu Gly
      290      295      300
Gly Thr Ala Gly Met Val Thr Leu Glu Asp Leu Val Glu Glu Ile Phe
      305      310      315      320
Gly Asp Ile Glu Asp Glu His Asp Thr Arg Lys Ile Ile Ala Lys Gln
      325      330      335
Leu Gly Pro His Thr Tyr Leu Val Ser Gly Arg Met Glu Ile Asp Asp
      340      345      350
Val Asn Glu Arg Phe Gly Leu Ser Leu Pro Glu Ser Asp Asp Tyr Leu
      355      360      365
Thr Val Ala Gly Phe Ile Leu Asn Ser His Gln Asn Ile Pro Gln Ala
      370      375      380
Asn Glu Val Val Glu Ile Ala Pro Tyr Thr Phe Thr Ile Leu Arg Ser
      385      390      395      400
Ser Ser Thr Lys Ile Glu Leu Val Lys Met Ser Ile Asp Asp Gln Ser
      405      410      415
Asn

```

## (2) INFORMATION FOR SEQ ID NO:25

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 672 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (iii) HYPOTHETICAL: YES

## (vi) ORIGINAL SOURCE:

- (A) ORGANISM: *Porphyromonas gingivalis*

## (ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION 1...672

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25

Met Tyr Ser Gly His His Lys Ile His Tyr Pro Phe Leu Ile Leu Leu

1	5	10	15
Val Cys Leu Ala Phe Ala Ala Cys Lys Ser Val Lys Leu Lys Asp Ala			
	20	25	30
Glu Lys Ala His Asp Arg Gln Glu Tyr Thr Lys Ala Ala Asp Met Tyr			
	35	40	45
Asn Thr Leu Tyr Arg Arg Thr Arg Arg Lys Gln Val Glu Met Lys Ala			
	50	55	60
Tyr Thr Ala Phe Arg Ser Gly Glu Asn Tyr Arg Ala Ala Gly Arg Gln			
	65	70	75
Ala Lys Ala Leu Arg Gly Tyr Leu Asn Ala Arg Arg Tyr Gly Tyr Pro			
	85	90	95
Asp Ser Val Val Leu Leu Arg Leu Ala Gln Thr Tyr Gln Gln Gly Gly			
	100	105	110
Asn Tyr Lys Glu Ala Glu Val Leu Phe Arg Gly Tyr Leu Glu Ala Tyr			
	115	120	125
Pro Lys Ser Tyr Phe Ala Ala Ile Gly Leu Glu Gly Cys Leu Phe Ala			
	130	135	140
Arg Gln Gln Lys Glu Tyr Pro Thr Arg Tyr Arg Ile Arg Arg Ala Ala			
	145	150	155
Glu Trp Asn Ser Ala Arg Gly Asp Phe Gly Pro Ala Tyr Ala Pro Asp			
	165	170	175
Ala Ser Ala Leu Tyr Phe Thr Ser Ser Arg Ser Lys Asp Asp Gly Leu			
	180	185	190
Asp Asn Ser Ser Ile Thr Gly Leu Lys Pro Asn Asp Ile Tyr Ile Ile			
	195	200	205
Lys Arg Asp Ala Gln Gly Arg Trp Gly Arg Pro Asp Ser Val Ser Gly			
	210	215	220
Gly Ile Asn Thr Pro Trp Asp Glu Gly Val Pro Thr Ile Thr Pro Asp			
	225	230	235
Gly Ser Thr Ile Tyr Tyr Thr Leu Ala Gln Gln Gly Ala Asp Tyr Asp			
	245	250	255
Arg Thr Val Gln Ile Tyr Ser Ala Ala Arg Ser Gly Glu Gly Gly Trp			
	260	265	270
Ser Asn Gly Ser Leu Val Asp Ile Met Arg Asp Ser Leu Arg Met Ala			
	275	280	285
Ala His Pro Ser Met Ser Ala Ser Gly Asp Tyr Leu Tyr Phe Val Ser			
	290	295	300
Asn Ile Gly Gly Ser Tyr Gly Gly Lys Asp Ile Tyr Arg Val Lys Val			
	305	310	315
Ser Asp Arg Ser Tyr Gly Ser Pro Glu Asn Leu Gly Pro Asp Ile Asn			
	325	330	335
Thr Pro Gly Asp Glu Met Phe Pro Phe Ile Asp Gly Asp Ser Thr Leu			
	340	345	350
Phe Phe Ala Ser Asp Gly His Ala Gly Leu Gly Gly Leu Asp Ile Phe			
	355	360	365
Lys Ala Thr Leu Asp Ser Thr Gly Gln Trp His Val Val Asn Met Gly			
	370	375	380
Gln Pro Val Asn Ser Ser Ala Asp Asp Phe Gly Leu Ala Val Glu Pro			
	385	390	395
Lys Gly Lys Asn Lys Glu Glu Ala Leu Pro Asp Asn Gly Val Lys Gly			
	405	410	415
Val Phe Cys Ser Asn Arg Gly Asp Ala Arg Gly Trp Pro His Leu Phe			
	420	425	430
His Phe Glu Leu Pro Ala Ile Tyr Thr Glu Ile Gln Gly Tyr Val Met			
	435	440	445
Asp Arg Glu Glu Asn Pro Ile Ala Gly Ala Thr Val Arg Ile Val Gly			
	450	455	460
Glu Arg Gly Pro Val Gly Gln Gly Phe Val Thr Thr Arg Asp Asp Gly			
	465	470	475
Ser Tyr Lys Met Ser Val Gln Gly Asp Thr Arg Tyr Val Met Leu Ala			
	485	490	495
Gly Ala Ser Gly Tyr Leu Asn Gln Tyr Val Glu Leu Lys Thr Asp Thr			
	500	505	510
Ala Lys Gln Ser Glu Thr Tyr Tyr Val Asp Phe Phe Leu Ala Ser Arg			

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 324 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

(A) ORGANISM: Porphyromonas gingivalis

```
(A) NAME/KEY: misc_feature
(B) LOCATION 1...324
```

Met 1	Cys	Leu	Glu	Pro 5	Ile	Ile	Ala	Pro	Ile 10	Ser	Ser	Glu	Leu	Leu 15	Glu
Gln	Glu	Leu	Thr 20	Ala	Asp	Arg	Phe	Leu 25	Arg	Met	Thr	Asn	Lys 30	Ala	Gly
Asn	Glu	Ile 35	Tyr	Val	Phe	Thr	Ala 40	Glu	Glu	Ala	Pro	His 45	Cys	Met	Lys
Glu	Val 50	Gly	Arg	Leu	Arg	Glu 55	Glu	Ala	Phe	Arg	His 60	Tyr	Gly	Gly	Gly
Thr 65	Gly	Lys	Ala	Ile 70	Asp	Ile	Asp	Glu	Phe 75	Asp	Thr	Met	Pro	Gly	Ser 80
Tyr	Lys	Gln	Leu	Ile 85	Val	Trp	Asp	Pro	Gln 90	Asn	Lys	Ala	Ile	Leu 95	Gly
Gly	Tyr	Arg	Phe 100	Ile	Tyr	Gly	Arg	Asp 105	Val	Ala	Phe	Asp 110	Thr	Asp	Gly
Lys	Pro	Leu 115	Leu	Ala	Thr	Ala	Glu 120	Met	Phe	Arg	Phe	Ser 125	Asp	Ala	Phe
Leu	His 130	Asp	Tyr	Leu	Pro	Tyr 135	Thr	Val	Glu	Leu	Gly 140	Arg	Ser	Phe	Val
Ser 145	Leu	Gln	Tyr	Gln	Ser 150	Thr	Arg	Met	Gly 155	Thr	Lys	Ala	Ile	Phe	Val 160
Leu	Asp	Asn	Leu	Trp 165	Asp	Gly	Ile	Gly	Ala 170	Leu	Thr	Val	Val	Asn 175	Pro
Glu	Ala	Leu	Tyr	Phe	Tyr	Gly	Lys	Val	Thr	Met	Tyr	Lys	Asp	Tyr	Asp

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 313 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: YES

(A) ORGANISM: Porphyromonas gingivalis

(A) NAME/KEY: misc\_feature  
(B) LOCATION 1...313

Met	Arg	Ala	Asn	Ile	Trp	Gln	Ile	Leu	Ser	Val	Ser	Val	Leu	Phe	Phe
1				5					10					15	
Phe	Gly	Thr	Ala	Ile	Gly	Gln	Ala	Gln	Ser	Arg	Asn	Arg	Thr	Tyr	Glu
			20					25					30		
Ala	Tyr	Val	Lys	Gln	Tyr	Ala	Asp	Glu	Ala	Ile	Arg	Gln	Met	Ser	Arg
			35				40					45			
Tyr	Asn	Ile	Pro	Ala	Ser	Ile	Thr	Ile	Ala	Gln	Ala	Leu	Val	Glu	Thr
	50					55					60				
Gly	Ala	Gly	Ala	Ser	Thr	Leu	Ala	Ser	Val	His	Asn	Asn	His	Phe	Gly
65					70					75				80	
Ile	Lys	Cys	His	Lys	Ser	Trp	Thr	Gly	Lys	Arg	Thr	Tyr	Arg	Thr	Asp
				85					90					95	
Asp	Ala	Pro	Asn	Glu	Cys	Phe	Arg	Ser	Tyr	Ser	Ala	Ala	Arg	Glu	Ser
			100					105					110		
Tyr	Glu	Asp	His	Ser	Arg	Phe	Leu	Leu	Gln	Pro	Arg	Tyr	Arg	Pro	Leu
			115				120					125			
Phe	Lys	Leu	Asp	Arg	Glu	Asp	Tyr	Arg	Gly	Trp	Ala	Thr	Gly	Leu	Gln
	130					135					140				
Arg	Cys	Gly	Tyr	Ala	Thr	Asn	Arg	Gly	Tyr	Ala	Asn	Leu	Leu	Ile	Lys
145					150					155				160	
Met	Val	Glu	Leu	Tyr	Glu	Leu	Tyr	Ala	Leu	Asp	Arg	Glu	Lys	Tyr	Pro
				165					170					175	
Ser	Trp	Phe	His	Lys	Ser	Tyr	Pro	Gly	Ser	Asn	Lys	Lys	Ser	His	Gln
			180					185					190		

